Group Sparse Kernelized Dictionary Learning for the Clustering of White Matter Fibers

Kuldeep Kumar and Christian Desrosiers

Ecole de technologie supérieure (ETS)
1100 Notre-Dame W., Montreal, Canada H3C1K3
kuldeep.kumar@etsmtl.ca, christian.desrosiers@etsmtl.ca

Abstract. This paper presents a novel method that combines kernelized dictionary learning and group sparsity to efficiently cluster white matter fiber tracts obtained from diffusion Magnetic Resonance Imaging (dMRI). Instead of having an explicit feature representation for the fibers, this method uses a non-linear kernel and specialized distance measures that can better learn complex bundles. Through the use of a global sparsity prior, the method also provides a soft assignment of fibers to bundles, making it more robust to overlapping fiber bundles and outliers. Furthermore, by using a group sparsity prior, it can automatically discard small and uninteresting bundles. We evaluate our method both qualitatively and quantitatively using expert labeled data, and compare it with state of the art approaches for this task.

1 Introduction

Due to its ability to infer the orientation of white matter fibers in-vivo and non-invasively, diffusion tensor imaging (DTI) has become an essential tool to study the microstructure of white matter in the brain. While extracting the individual fiber tracts from DTI data, a process known as tractography, is important to visualize the connection pathways in the brain, this process typically produces a large number of tracts which makes their analysis complex. To facilitate this analysis, it is often necessary to group the individual tracts into larger clusters, called bundles.

Methods proposed for the fiber clustering problem can be categorized in terms of the features and distance measures used to group the fibers into bundles. Features proposed to represent fibers include the distribution parameters (mean and covariance) of points along the fiber [2] and B-splines [11]. Approaches using such explicit features typically suffer from two problems: they are sensitive to the length and endpoint positions of the fibers and/or are unable to capture their full shape. Instead of using explicit features, fibers can also be compared using specialized distance measures. Popular distance measures for this task include the Hausdorff distance, the Minimum Direct Flip (MDF) distance and the Mean Closest Points (MCP) distance [3, 12]. Fiber clustering approaches can also be divided with respect to the clustering methods used, which include manifold embedding based approaches like spectral clustering and normalized cuts [2], agglomerative approaches like hierarchical clustering [3], k-means, and k-nearest

neighbors [12]. Several studies have also focused on incorporating anatomical features into the clustering [14] and on clustering large multi-subject datasets [9].

Recently, several researchers have studied the connection between clustering and factorization problems like dictionary learning [15] and non-negative matrix factorization [10]. For instance, dictionary learning has been shown to be a generalization of the traditional clustering problem, in which objects can be assigned to more than one cluster. In fiber clustering, such soft assignments are desirable since fiber bundles often overlap each other. Using a soft clustering, instead of hard one, can also make the method more robust to outliers (e.g., false fibers generated during tractography) that do not belong to any real cluster. Moreover, researchers have also recognized the advantages of applying kernels to existing clustering methods, like the k-means algorithm [4], as well as to dictionary learning approaches [13]. Such "kernelized" methods better capture the non-linear relations in the data.

The major contribution of this paper is a novel fiber clustering approach based on kernelized dictionary learning. By modeling the fiber clustering task as a dictionary learning problem, this approach provides a soft assignment of fibers to bundles, which makes it more robust to overlapping bundles and outliers. Furthermore, through the use of a non-linear kernel, it avoids the need to specify explicit features for the fibers, and can facilitate the separation of clusters in a manifold space. Also, by having both a global and group sparsity prior, our approach can control the minimum membership of fibers to bundles as well as the size of these bundles. This makes it more robust to the selection of the number of clusters in the output, a parameter which can be hard to tune, and allows it to automatically discard insignificant clusters. To our knowledge, this work is the first to combine group sparsity and kernelized dictionary learning. Our results on the fiber clustering problem show the potential of this approach for other medical imaging applications.

2 The proposed approach

2.1 The clustering problem

Before presenting our proposed approach, we first define the clustering problem and underline its link to dictionary learning. Let $\boldsymbol{X} \in \mathbb{R}^{d \times n}$ be the data matrix of n fibers, where each column contains the feature vector $\boldsymbol{x}_i \in \mathbb{R}^d$ of a fiber tract i. The traditional (hard) clustering problem can be defined as assigning each fiber to a bundle from a set of k bundles, such that fibers are as close as possible to their assigned bundle's prototype (i.e., cluster center). Let $\Psi^{k \times n}$ be the set of all $k \times n$ cluster assignment matrices (i.e., matrices in which each row has a single non-zero value equal to one), this problem can be expressed as finding the matrix $\boldsymbol{D}$ of k bundle prototypes and the fiber-to-bundle assignment matrix $\boldsymbol{W}$ that minimize the following cost function:

$$\min_{\substack{\boldsymbol{D} \in \mathbb{R}^{d \times k} \\ \boldsymbol{W} \in \Psi^{k \times n}}} \quad \frac{1}{2} ||\boldsymbol{X} - \boldsymbol{D}\boldsymbol{W}||_F^2. \tag{1}$$

This formulation of the clustering problem can be seen as a special case of dictionary learning, where $\boldsymbol{D}$ is the dictionary and $\boldsymbol{W}$ is constrained to be a cluster assignment matrix, instead of constraining its sparsity.

While solving this clustering problem is NP-hard, optimizing $\boldsymbol{W}$ or $\boldsymbol{D}$ individually is easy. Thus, for a given dictionary $\boldsymbol{D}$, the optimal $\boldsymbol{W}$ assigns each fiber i to the prototype k closest to its feature vector:

$$w_{ki} \;=\; \begin{cases} 1 : \text{if } k = \arg\min_{k'} \; ||\boldsymbol{x}_i - \boldsymbol{d}_{k'}||^2, \\ 0 : \text{otherwise.} \end{cases} \tag{2}$$

Likewise, for a fixed $\boldsymbol{W}$, the optimal dictionary is found by solving a simple linear regression problem:

$$\boldsymbol{D} \;=\; \boldsymbol{X}\boldsymbol{W}^\top\big(\boldsymbol{W}\boldsymbol{W}^\top\big)^{-1}. \tag{3}$$

This suggest the following heuristic: starting with a dictionary containing a random subset of the columns of $\boldsymbol{X}$, optimize $\boldsymbol{D}$ and $\boldsymbol{W}$ alternatively, until convergence.

This clustering problem and simple heuristic correspond to the well-known k-means algorithm. With respect to dictionary learning, the dictionary update step described above is known as the Method of Optimal Directions (MOD) [1]. Although k-SVD [1] could also be used for this task, this technique focuses on learning large dictionaries efficiently and sacrifices the optimality of the dictionary update step to do so. In our case, the dictionary size corresponds to the number k of bundles (i.e., *clusters*), which is quite small in comparison to the number of tracts. Thus, updating the dictionary using MOD is quite fast.

2.2 Group sparse kernel dictionary learning

The k-means approach described in the previous section suffers from four important problems. First, it requires to encode fibers as a set of features, which is problematic due to the variation in their length and endpoints. Second, it assumes linear relations between the fibers and bundle prototypes, while these relations could be better defined in a non-linear subspace (i.e., the manifold). Third, it performs a hard clustering of the fibers, which can lead to poor results in the presence of overlapping bundles and outliers. Finally, it may find insignificant bundles (e.g., bundles containing only a few fibers) when the parameter controlling the number of clusters is not properly set.

To overcome these problem, we present a new clustering method based on group sparse kernelized dictionary learning. Let $\phi : \mathbb{R}^d \to \mathbb{R}^q$ be a fiber mapping function such that $k(\boldsymbol{x}_i, \boldsymbol{x}_j) \;=\; \phi(\boldsymbol{x}_i)^\top\phi(\boldsymbol{x}_j)$ corresponds to a similarity kernel. Moreover, denote by $\boldsymbol{\Phi}$ the matrix of mapped fiber tracts, i.e., $\boldsymbol{\Phi} = \phi(\boldsymbol{X})$, and let $\boldsymbol{K} = \boldsymbol{\Phi}^\top\boldsymbol{\Phi}$ be the kernel matrix. We reformulate the clustering problem as finding the dictionary $\boldsymbol{D}$ and non-negative weight matrix $\boldsymbol{W}$ minimizing the following problem:

$$\min_{\substack{\boldsymbol{D}\in\mathbb{R}^{q\times k} \\ \boldsymbol{W}\in\mathbb{R}_+^{k\times n}}} \; f(\boldsymbol{D},\boldsymbol{W}) \;=\; \frac{1}{2}||\boldsymbol{\Phi}-\boldsymbol{D}\boldsymbol{W}||_F^2 + \lambda_1||\boldsymbol{W}||_1 + \lambda_2||\boldsymbol{W}||_{2,1} + \frac{\lambda_3}{2}||\boldsymbol{D}||_F^2. \tag{4}$$

In this formulation, $||\boldsymbol{W}||_1 = \sum_{i=1}^{K}\sum_{j=1}^{N}|w_{ij}|$ is an L_1 norm prior which enforces global sparsity of $\boldsymbol{W}$, and $||\boldsymbol{W}||_{2,1} = \sum_{i=1}^{K}||\boldsymbol{w}_{i\cdot}||_2$ is a mixed $L_{2,1}$ norm prior imposing the vector of row norms to be sparse. Concretely, the L_1 norm prior limits the "membership" of fibers to a small number of bundles, while the $L_{2,1}$ prior penalizes the clusters containing only a few fibers. The Forbenius norm prior on $\boldsymbol{D}$ is used to avoid numerical problems when $\boldsymbol{W}$ is singular (i.e., when one or more clusters are empty). Parameters $\lambda_1, \lambda_2, \lambda_3 \geq 0$ control the trade-off between these three properties and the reconstruction error (i.e., the first term of the cost function).

Using an optimization approach similar to k-means, we alternate between updating the dictionary $\boldsymbol{D}$ and the weight matrix $\boldsymbol{W}$. Since the dictionary prototypes are defined in the kernel space, $\boldsymbol{D}$ cannot be computed explicitly. To overcome this problem, we follow the strategy proposed in [13] and define the dictionary as $\boldsymbol{D} = \boldsymbol{\Phi}\boldsymbol{A}$, where $\boldsymbol{A} \in \mathbb{R}^{n\times k}$. Using this formulation, $\boldsymbol{A}$ can be computed as follows:

$$\boldsymbol{A} \;=\; \boldsymbol{W}^\top\big(\boldsymbol{W}\boldsymbol{W}^\top + \lambda_3\boldsymbol{I}\big)^{-1}. \tag{5}$$

Matrix $\boldsymbol{A}$ is initialized as a random selection matrix (i.e., random subset of columns in the identity matrix), which is equivalent to using a random subset of the transformed fibers (i.e., subset of columns in $\boldsymbol{\Phi}$) as the initial dictionary.

To update $\boldsymbol{W}$, we use an Alternating Direction Method of Multipliers (ADMM) method. First, we separate the problem in two sub-problems, one considering only the reconstruction error and the second considering only the (group) sparsity terms and non-negativity constraints, by introducing ancillary matrix $\boldsymbol{Z}$. The problem can then be reformulated as follows:

$$\min_{\substack{\boldsymbol{W}\in\mathbb{R}^{k\times n}\\ \boldsymbol{Z}\in\mathbb{R}_+^{k\times n}}} \;\frac{1}{2}||\boldsymbol{\Phi}-\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{W}||_F^2 \;+\; \lambda_1||\boldsymbol{Z}||_1 \;+\; \lambda_2||\boldsymbol{Z}||_{2,1}, \quad \text{s.t. } \boldsymbol{W}=\boldsymbol{Z}. \tag{6}$$

We then convert this constrained problem using an Augmented Lagrangian formulation with multipliers $\boldsymbol{U}$:

$$\min_{\substack{\boldsymbol{W},\boldsymbol{U}\in\mathbb{R}^{k\times n}\\ \boldsymbol{Z}\in\mathbb{R}_+^{k\times n}}} \;\frac{1}{2}||\boldsymbol{\Phi}-\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{W}||_F^2 \;+\; \lambda_1||\boldsymbol{Z}||_1 \;+\; \lambda_2||\boldsymbol{Z}||_{2,1} \;+\; \frac{\mu}{2}||\boldsymbol{W}-\boldsymbol{Z}+\boldsymbol{U}||_F^2. \tag{7}$$

In an inner loop, we update $\boldsymbol{W}$, $\boldsymbol{Z}$ and $\boldsymbol{U}$ alternatively, until convergence (i.e., $||\boldsymbol{W}-\boldsymbol{Z}||_F^2$ is below some threshold). To update $\boldsymbol{W}$, we derive the objective function with respect to this matrix and set the result to 0, yielding:

$$\boldsymbol{W} \;=\; \big(\boldsymbol{A}^\top\boldsymbol{K}\boldsymbol{A}+\mu\boldsymbol{I}\big)^{-1}\big(\boldsymbol{A}^\top\boldsymbol{K}+\mu(\boldsymbol{Z}-\boldsymbol{U})\big). \tag{8}$$

Optimizing $\boldsymbol{Z}$ corresponds to solving a group sparse proximal problem (see [7]). This can be done in two steps. First, we do a L_1-norm shrinkage by applying the non-negative soft-thresholding operator to each element of $\boldsymbol{W}+\boldsymbol{U}$:

$$\hat{z}_{ij} \;=\; S^{+}_{\frac{\lambda_1}{\mu}}\big(w_{ij}+u_{ij}\big) \;=\; \max\Big\{w_{ij}+u_{ij}-\frac{\lambda_1}{\mu},\,0\Big\}, \quad i\leq K,\; j\leq N. \tag{9}$$

Then, $\boldsymbol{Z}$ is obtained by applying a group shrinkage on each row of $\hat{\boldsymbol{Z}}$:

$$\boldsymbol{z}_{i\boldsymbol{\cdot}} \;=\; \max\Big\{||\hat{\boldsymbol{z}}_{i\boldsymbol{\cdot}}||_2 - \frac{\lambda_2}{\mu},\, 0\Big\} \cdot \frac{\hat{\boldsymbol{z}}_{i\boldsymbol{\cdot}}}{||\hat{\boldsymbol{z}}_{i\boldsymbol{\cdot}}||_2}, \quad i \leq K. \tag{10}$$

Finally, as in standard ADMM methods, the Lagrangian multipliers are updated as follows:

$$\boldsymbol{U}' \;=\; \boldsymbol{U} \;+\; (\boldsymbol{W} - \boldsymbol{Z}). \tag{11}$$

2.3 Algorithm summary and complexity

The clustering process of our proposed method is summarized in Algorithm 1. In this algorithm, the user provides a matrix $\boldsymbol{Q}$ of pairwise fiber distances (see Section 3 for more details), the maximum number of clusters k, as well as the trade-off parameters $\lambda_1, \lambda_2, \lambda_3$, and obtains as output the dictionary matrix $\boldsymbol{A}$ and the cluster assignment weights $\boldsymbol{W}$. At each iteration, $\boldsymbol{W}$, $\boldsymbol{Z}$ and $\boldsymbol{U}$ are updated by running at most T_{in} ADMM loops, and are then used to update $\boldsymbol{A}$. This process is repeated until T_{out} iterations have been completed or the cost function $f(\boldsymbol{D}, \boldsymbol{W})$ converged. The soft assignment of $\boldsymbol{W}$ can be converted to a hard clustering by assigning each fiber i to the bundle k for which w_{ik} is maximum.

The complexity of this algorithm is mainly determined by the initial kernel computation, which takes $O(n^2)$ operations, and updating the assignment weights in each ADMM loop, which has a total complexity in $O(T_{\text{out}} \cdot T_{\text{in}} \cdot k^2 \cdot n)$. Since T_{out}, T_{in} and k are typically much smaller than n, the main bottleneck of the method lies in computing the pairwise distances $\boldsymbol{Q}$ used as input. For datasets having a large number of fibers (e.g., more than $n = 100,000$ fibers), this matrix could be computed using an approximation strategy such as the the Nyström method [6].

3 Experiments

We evaluated the performance of our proposed method on a dataset of expert labeled bundles, provided by the Sherbrooke Connectivity Imaging Laboratory (SCIL). The source dMRI data was acquired from a 25 year old healthy right-handed volunteer and is described in [5]. We used 10 of the largest bundles, consisting of 4449 fibers identified from the cingulum, corticospinal tract, superior cerebellar penduncle and other prominent regions. Figure 2(b) shows the coronal and sagittal plane view of the ground truth set.

Although our method has several parameters, only two of them require data specific tuning: λ_1 and λ_2. The RBF kernel parameter γ depends on the distance measure used, not the dataset. For these experiments, we used the Mean Closest Points (**MCP**) distance [3] to compute the pairwise fiber distances $\boldsymbol{Q}$, and set γ to 0.01. Also, λ_3 and μ correspond to regularization parameters and should be set to a small positive value. In our experiments, we have used $\lambda_3 = 10^{-6}$ and $\mu = 0.01$ for these parameters. According to Eq. 9, λ_1/μ corresponds to a minimum threshold for the assignment weights. As shown in Figure 1(a), this

Algorithm 1: ADMM method for group sparse kernelized clustering

Input: Pairwise fiber distance matrix $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$;
Input: The maximum number of fiber bundles k;
Input: The RBF kernel parameter γ;
Input: The cost trade-off parameters $\lambda_1, \lambda_2, \lambda_3$ and Lagrangian parameter μ;
Input: The maximum number of inner and outer loop iterations $T_{\text{in}}, T_{\text{out}}$;
Output: The dictionary $\boldsymbol{A} \in \mathbb{R}^{n \times k}$ and assignment weights $\boldsymbol{W} \in \mathbb{R}_{+}^{n \times k}$;

Initialize the kernel matrix: $k_{ij} = \exp(-\gamma \cdot q_{ij}^2)$;
Initialize $\boldsymbol{A}$ as a random selection matrix and t_{out} to 0;

while $f(\boldsymbol{D}, \boldsymbol{W})$ *not converged and* $t_{\text{out}} \leq T_{\text{out}}$ **do**

  Initialize $\boldsymbol{U}$ and $\boldsymbol{Z}$ to all zeros and t_{in} to zero;

  while $||\boldsymbol{W} - \boldsymbol{Z}||_F^2$ *not converged and* $t_{\text{in}} \leq T_{\text{in}}$ **do**

    Update $\boldsymbol{W}$, $\boldsymbol{Z}$ and $\boldsymbol{U}$:

$$\boldsymbol{W} \leftarrow \big(\boldsymbol{A}^\top \boldsymbol{K} \boldsymbol{A} + \mu \boldsymbol{I}\big)^{-1} \big(\boldsymbol{A}^\top \boldsymbol{K} + \mu(\boldsymbol{Z} - \boldsymbol{U})\big);$$

$$\hat{z}_{ij} \leftarrow \max\Big\{w_{ij} + u_{ij} - \frac{\lambda_1}{\mu}, 0\Big\}, \quad i \leq K,\ j \leq N;$$

$$\boldsymbol{z}_{i\bullet} \leftarrow \max\Big\{||\hat{\boldsymbol{z}}_{i\bullet}||_2 - \frac{\lambda_2}{\mu}, 0\Big\} \cdot \frac{\hat{\boldsymbol{z}}_{i\bullet}}{||\hat{\boldsymbol{z}}_{i\bullet}||_2}, \quad i \leq K;$$

$$\boldsymbol{U} \leftarrow \boldsymbol{U} + \big(\boldsymbol{W} - \boldsymbol{Z}\big);$$

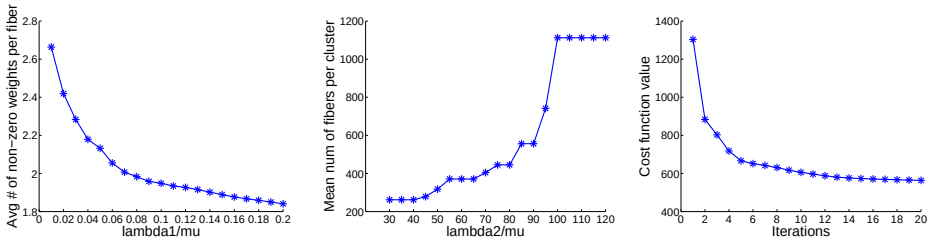
    $t_{\text{in}} \leftarrow t_{\text{in}} + 1$;

  Update dictionary: $\boldsymbol{A} \leftarrow \boldsymbol{W}^\top \big(\boldsymbol{W}\boldsymbol{W}^\top + \lambda_3 \boldsymbol{I}\big)^{-1}$;
  $t_{\text{out}} \leftarrow t_{\text{out}} + 1$;

return $\{\boldsymbol{A}, \boldsymbol{W}\}$;

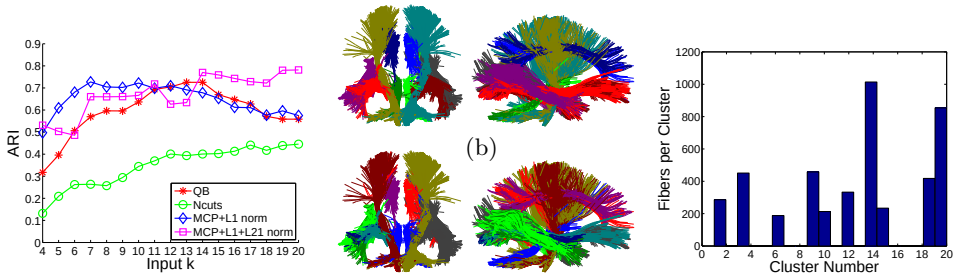
value can be used to control the mean number of non-zero weights per fiber (i.e., how soft or hard is the clustering). Likewise, λ_2/μ is a minimum threshold on the total membership to a bundle and, as shown in Figure 1(b), controls the size of bundles in the output. Finally, following the convergence rate shown in Figure 1(c), we have used $T_{\text{out}} = 20$ for the maximum number of iterations. The same value was used for the number of inner loop iterations (i.e., $T_{\text{in}} = 20$).

Figure 2(a) shows the mean Adjusted Rand Index (**ARI**) [12] obtained by our method, denoted by **MCP+L1+L21**, over 5 runs with different random initializations. We compared this method with two well-known fiber clustering approaches: QuickBundles (**QB**) [8] and Normalized cuts (**Ncuts**) [2]. QuickBundles recursively groups fibers between which the Minimum Direct Flip (**MDF**) distance is below a specified threshold. Ncuts performs a spectral embedding of the fibers encoded as the mean and covariance parameters of the points distribution, and then clusters the embedded fibers using a recursive partitioning strategy or k-means. Based on earlier results, we used 25 eigenvectors for the embedding and k-means for clustering. We also tested our method without the group sparsity prior (i.e., using $\lambda_2 = 0$) and called **MCP+L1** this simplified model.



(a) (b) (c)

Fig. 1: **(a)** Mean number of non-zero assignment weights per fiber, for $\lambda_2/\mu = 80$ and increasing λ_1/μ. **(b)** Mean number of fibers per bundle, for $\lambda_1/\mu = 0.1$ and increasing λ_2/μ. **(c)** Cost function value at each iteration of a sample run.



(a) (c) (d)

Fig. 2: **(a)** Mean ARI of QB, Ncuts, MCP+L1 ($\lambda_1/\mu = 0.1$, $\lambda_2/\mu = 0$) and MCP+L1+L21 ($\lambda_1/\mu = 0.1$, $\lambda_2/\mu = 80$), for increasing k. **(b)-(c)** Ground truth bundles and clustering output of MCP+L1+L21 for $k = 20$. **(d)** Distribution of bundle sizes corresponding to this output.

From these results, we see that Ncuts performs worse than all other methods. This is possibly due to the fact that the features used to encode the fibers do not fully capture their shape. Moreover, we observe that the peak ARI of QuickBundles is similar to that of MCP+L1, but the latter peaks closer to the true number of bundles (i.e., 10). Finally, we see that the MCP+L1+L21 method, which also considers group sparsity, obtains the highest ARI and is less sensitive to the value of k given as input. The bundles obtained by this method for $k = 20$ are presented in Figure 2(c). As shown in Figure 2(d), this clustering contains the same number of clusters as the ground truth, even though the maximum number of clusters was set to $k = 20$.

4 Conclusion

We have presented a new fiber clustering approach based on dictionary learning. This approach uses a non-linear kernel which avoids having to define features for the fibers and can represent complex bundles. Furthermore, by using an L_1 norm prior, instead of hard clustering constraints, it is more robust to overlap-

ping bundles and outliers. Finally, since it also includes a group sparsity prior, our approach can find more interesting bundles than other methods for this task. Experiments conducted on expert labeled data show our methods to outperform state of the art fiber clustering approaches such as QuickBundles and Normalized Cuts. In future work, we will extend the proposed model to incorporate anatomical information in the form of atlases.

References

1. Aharon, M., Elad, M., Bruckstein, A.: k-svd: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing 54(11), 4311–4322 (2006)
2. Brun, A., Knutsson, H., et al.: Clustering fiber traces using normalized cuts. In: MICCAI 2004, pp. 368–375. Springer (2004)
3. Corouge, I., Gouttard, S., Gerig, G.: Towards a shape model of white matter fiber bundles using diffusion tensor mri. In: IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004. pp. 344–347. IEEE (2004)
4. Dhillon, I., Guan, Y., Kulis, B.: Kernel k-means: spectral clustering and normalized cuts. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 551–556. ACM (2004)
5. Fortin, D., Aubin-Lemay, C., et al.: Tractography in the study of the human brain: a neurosurgical perspective. The Canadian Journal of Neurological Sciences 39(6), 747–756 (2012)
6. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the nystrom method. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(2), 214–225 (2004)
7. Friedman, J., Hastie, T., Tibshirani, R.: A note on the group lasso and a sparse group lasso. arXiv preprint arXiv:1001.0736 (2010)
8. Garyfallidis, E., Brett, M., et al.: Quickbundles, a method for tractography simplification. Frontiers in neuroscience 6 (2012)
9. Guevara, P., Duclap, D., et al.: Automatic fiber bundle segmentation in massive tractography datasets using a multi-subject bundle atlas. Neuroimage 61(4), 1083–1099 (2012)
10. Kim, H., Park, H.: Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. Bioinformatics 23(12), 1495–1502 (2007)
11. Maddah, M., Crimson, W., Warfield, S.: Statistical modeling and em clustering of white matter fiber tracts. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006. pp. 53–56. IEEE (2006)
12. Moberts, B., Vilanova, A., van Wijk, J.: Evaluation of fiber clustering methods for diffusion tensor imaging. In: Visualization, 2005. pp. 65–72. IEEE (2005)
13. Nguyen, H., Patel, V., Nasrabadi, N., Chellappa, R.: Kernel dictionary learning. In: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on. pp. 2021–2024. IEEE (2012)
14. O'Donnell, L., Westin, C.F.: Automatic tractography segmentation using a high-dimensional white matter atlas. IEEE Transactions on Medical Imaging 26(11), 1562–1575 (2007)
15. Sprechmann, P., Sapiro, G.: Dictionary learning and sparse coding for unsupervised clustering. In: ICASSP 2010. pp. 2042–2045. IEEE (2010)